Evaluation und Metaanalyse

MSc Klinische Psychologie und Psychotherapie

SoSe 2024

Prof. Dr. Dirk Ostwald

(5) Linear Mixed Models Formulierung

Überblick

Motivation

- Linear Mixed Models sind eine Weiterentwicklung des klassischen Allgemeinen Linearen Modells.
- Durch iterative Schätzverfahren sind Linear Mixed Models in den letzten 50 Jahren populär geworden.
- In **R** sind Linear Mixed Models im Sinne der Pakete `nlme` und `lme4` sehr populär.
- Klassische Anwendungen von Linear Mixed Models sind Mehrebenen- und Longitudinalanalysen.
- Viele Modelle sind Spezialfälle von Linear Mixed Models, z.B. Bayesianische ALM Schätzung.
- Fixed- und Random-Effects Modelle der Metaanalyse sind spezielle Linear Mixed Models.
- Die Restricted Maximum-Likelihood-Schätzung (ReML) ist eng mit Linear Mixed Models verwoben.
- Linear Mixed Models sind State-of-the-Art statistische Inferenzmodelle in vielen Domänen.

Wir erinnern zunächst an das Allgemeine Lineare Modell und formulieren dann ein Linear Mixed Model.

Mit der Generalized-Least-Squares und der ReML Schätzung führen wir dann Modellschätzverfahren ein.

Wir betrachten schließlich Anwendungsbeispiele

(1) Anwendung von Linear Mixed Models im Bereich Metaanalyse
(2) Anwendung von Linear Mixed Models im Bereich Veränderungsmessung bei Psychotherapie
(3) Anwendung von Linear Mixed Models im Bereich natürlich geclusterter Daten

Allgemeines Lineares Modell

Linear Mixed Models

Selbstkontrollfragen

Allgemeines Lineares Modell

Linear Mixed Models

Selbstkontrollfragen

Integrativer Rahmen für viele inferenzstatistische Standardverfahren

- T-Tests
- Einfache und multiple Regression
- Einfaktorielle und mehrfaktorielle Varianzanalyse
- Kovarianzanalyse

Matrizen und Multivariate Normalverteilungen

Aktuelle Kursiteration im 2. FS Psychologie

Definition (Allgemeines Lineares Modell)

Es sei

$$y = X\beta + \varepsilon, \tag{1}$$

wobei

- y ein n-dimensionaler beobachtbarer Zufallsvektor ist, der *Daten* genannt wird,
- $X \in \mathbb{R}^{n \times p}$ mit $n > p$ eine vorgegebene Matrix ist, die *Designmatrix* genannt wird,
- $\beta \in \mathbb{R}^p$ ein unbekannter Parametervektor ist, der *Betaparametervektor* genannt wird,
- ε ein n-dimensionaler nicht-beobachtbarer Zufallsvektor ist, der *Zufallsfehler* genannt wird und für den angenommen wird, dass mit einem unbekannten Varianzparameter $\sigma^2 > 0$ gilt, dass

$$\varepsilon \sim N\left(0_n, \sigma^2 I_n\right). \tag{2}$$

Dann heißt (1) *Allgemeines Lineares Modell (ALM)*.

Allgemeines Lineares Modell

Bemerkungen

- Wir nehmen durchgängig an, dass $X \in \mathbb{R}^{n \times p}$ vollen Spaltenrang hat, also dass $\text{rg}(X) = p$.
- y ist ein Zufallsvektor, weil er aus der Addition des Zufallsvektors ε zu dem Vektor $X\beta \in \mathbb{R}^n$ resultiert.
- Wir nennen $X\beta \in \mathbb{R}^n$ den *deterministischen Modellaspekt* und ε den *probabilistischen Modellaspekt*.
- $n \in \mathbb{N}$ bezeichnet durchgängig die Anzahl an Datenpunkten.
- $p \in \mathbb{N}$ bezeichnet durchgängig die Anzahl an Betaparametern.
- Die Gesamtzahl an Parametern des ALMs ist $p + 1$ (p Betaparameterkomponenten und 1 Varianzparameter).
- Der Betaparametervektor wird auch *Gewichtsvektor* oder *Effektvektor* genannt.
- Weil der Kovarianzmatrixparameter von ε als sphärisch angenommen wird, sind die $\varepsilon_1, ..., \varepsilon_n$ unabhängige normalverteilte Zufallsvariablen mit identischem Varianzparameter; weil zusätzlich der Erwartungswertparameter von ε als 0_n angenommen wird, sind die $\varepsilon_1, ..., \varepsilon_n$ auch identisch normalverteilte Zufallsvariablen.
- Für jede Komponente $y_i, i = 1, ..., n$ von y impliziert (1) nach Definition des Matrixprodukts, dass

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{ip}\beta_p + \varepsilon_i \text{ mit } \varepsilon_i \sim N\left(0, \sigma^2\right), \tag{3}$$

wobei $x_{ij} \in \mathbb{R}$ das ijte Element der Designmatrix X bezeichnet.

Theorem (Parameterschätzer des Allgemeinen Linearen Modells)

Es sei

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n) \tag{4}$$

das ALM. Dann gilt, dass der *Betaparameterschätzer*

$$\hat{\beta} := \left(X^T X\right)^{-1} X^T y. \tag{5}$$

die Summe der Abweichungsquadrate minimiert,

$$\hat{\beta} = \arg\min_{\tilde{\beta}} (y - X\tilde{\beta})^T (y - X\tilde{\beta}), \tag{6}$$

und dass $\hat{\beta}$ ein unverzerrter Maximum-Likelihood Schätzer von $\beta \in \mathbb{R}^p$ ist. Weiterhin gilt, dass der *Varianzparameterschätzer*

$$\hat{\sigma}^2 := \frac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{n - p} \tag{7}$$

ein unverzerrter Schätzer von $\sigma^2 > 0$ ist.

Bemerkungen

- Wir verzichten auf einen Beweis.
- $\hat{\beta}$ hängt von den Daten und der Designmatrix ab.
- $\hat{\sigma}^2$ hängt von den Daten, der Designmatrix und dem Betaparameterschätzer ab.

Theorem (T-Statistik)

Es sei

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n) \tag{8}$$

das ALM. Weiterhin seien

$$\hat{\beta} := (X^T X)^{-1} X^T y \text{ und } \hat{\sigma}^2 := \frac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{n-p} \tag{9}$$

die Betaparameter- und Varianzparameterschätzer, respektive. Schließlich sei für einen Kontrastgewichtsvektor $c \in \mathbb{R}^p$ und einen Parameter $\beta_0 \in \mathbb{R}^p$

$$T := \frac{c^T \hat{\beta} - c^T \beta_0}{\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} c}} \tag{10}$$

die T-Statistik. Dann gilt

$$T \sim t(\delta, n-p) \text{ mit } \delta := \frac{c^T \beta - c^T \beta_0}{\sqrt{\sigma^2 c^T (X^T X)^{-1} c}}. \tag{11}$$

Bemerkungen

- Wir verzichten auf einen Beweis.
- T ist eine Funktion der Parameterschätzer, δ ist eine Funktion der wahren, aber unbekannten, Parameter
- Für $c^T \beta = c^T \beta_0$, also bei Zutreffen der Nullhypothese, gilt $\delta = 0$ und damit $T \sim t(n-p)$.
- Für $c^T \beta \neq c^T \beta_0$ kann die Verteilung von T zur Herleitung von Powerfunktionen benutzt werden.

Theorem (Konfidenzintervalle für Betaparameterkomponenten)

Es sei

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n) \tag{12}$$

das ALM, $\hat{\beta}$ und $\hat{\sigma}^2$ seien die Betaparameter- und Varianzparameterschätzer, respektive und und für ein $\delta \in]0, 1[$ sei

$$t_\delta := \Psi^{-1}\left(\frac{1+\delta}{2}; n-p\right). \tag{13}$$

Schließlich sei für $j = 1, ...p$

$$\lambda_j := \left(\left(X^T X\right)^{-1}\right)_{jj} \text{ das } j\text{te Diagonalelement von } \left(X^T X\right)^{-1}. \tag{14}$$

Dann ist für $j = 1, ...p$

$$\kappa_j := \left[\hat{\beta}_j - \hat{\sigma}\sqrt{\lambda_j} t_\delta, \hat{\beta}_j + \hat{\sigma}\sqrt{\lambda_j} t_\delta\right] \tag{15}$$

ein δ-Konfidenzintervall für die $j$te Komponente β_j des Betaparameters $\beta = (\beta_1, ..., \beta_p)^T$.

Bemerkungen

- Ψ^{-1} bezeichnet die inverse kumulative Verteilungsfunktion der t-Verteilung
- Intuitiv gilt im Vergleich zum Konfidenzintervall für den Erwartungswertparameter bei Normalverteilung

$$\hat{\beta}_j \approx \bar{y}, \hat{\sigma} \approx S, \sqrt{\lambda_j} \approx \sqrt{n^{-1}} \text{ und } t_\delta = t_\delta, \tag{16}$$

vgl. (11) Konfidenzintervalle in Wahrscheinlichkeitstheorie und Frequentistische Inferenz.

Zweistichproben-T-Test Szenario

- Zwei Gruppen (Stichproben) randomisierter experimenteller Einheiten.
- Annahme unabhängiger identischer Normalverteilungen $N(\mu_1, \sigma^2)$ und $N(\mu_2, \sigma^2)$.
- μ_1, μ_2 und σ^2 unbekannt.
- Annahme eines identischen Varianzparameters für beide Gruppen.
- Quantifizieren der Unsicherheit beim inferentiellen Vergleich von μ_1 mit μ_2 beabsichtigt.
- Beispiel : Zwei Gruppen Design mit Treatment- und Kontrollgruppe.

Allgemeines Lineares Modell | Zweistichproben-T-Tests

Beispieldatensatz (T $=$ Treatment, C $=$ Control, $n_t = 15$ Patient:innen, $n_c = 15$ Patient:innen)

ID	Group	PreBDI	PosBDI	dBDI
1	T	28	31	-3
2	T	31	27	4
3	T	27	28	-1
4	T	35	27	8
5	T	31	23	8
6	T	27	26	1
7	T	32	26	6
8	T	32	27	5
9	T	32	30	2
10	T	29	29	0
11	T	35	26	9
12	T	31	26	5
13	T	28	29	-1
14	T	23	29	-6
15	T	34	25	9
16	C	30	28	2
17	C	30	31	-1
18	C	33	32	1
19	C	33	30	3
20	C	32	33	-1
21	C	33	31	2
22	C	32	28	4
23	C	30	31	-1
24	C	24	26	-2
25	C	32	35	-3
26	C	30	36	-6
27	C	30	29	1
28	C	25	27	-2
29	C	28	32	-4
30	C	31	30	1

Definition (Zweistichproben-T-Test-Modell)

y_{ij} mit $i = 1, 2$ und $j = 1, ..., n_i$ seien Zufallsvariablen, die die Datenpunkte eines Zweistichproben-T-Test Anwendungsszenarios modellieren. Dann hat das *Zweistichproben-T-Test-Modell* die strukturelle Form

$$y_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, 2, j = 1, ..., n_i \text{ mit } \mu_i \in \mathbb{R} \text{ und } \sigma^2 > 0, \tag{17}$$

die Datenverteilungsform

$$y_{ij} \sim N(\mu_i, \sigma^2) \text{ u.i.v. für } i = 1, 2, j = 1, ..., n_i \text{ mit } \mu_i \in \mathbb{R} \text{ und } \sigma^2 > 0, \tag{18}$$

und für den Datenvektor $y = (y_{11}, ..., y_{1n_1}, y_{21}, ..., y_{2n_2})^T$ und $n := n_1 + n_2$ die Designmatrixform

$$y = X\beta + \varepsilon \text{ mit } X := \begin{pmatrix} 1_{n_1} & 0_{n_1} \\ 0_{n_2} & 1_{n_2} \end{pmatrix} \in \mathbb{R}^{n \times 2}, \beta := \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \in \mathbb{R}^2, \varepsilon \sim N(0_n, \sigma^2 I_n), \sigma^2 > 0. \tag{19}$$

Bemerkungen

- i indiziert die Gruppen, j indiziert die Daten in jeder Gruppe.
- n_1 und n_2 repräsentieren die Gruppengrößen, n repräsentiert die Gesamtanzahl an Datenpunkten.
- Die Anzahl der Betaparameter ist $p = 2$.

Datensimulation

```
# Libraries
library(MASS)                                 # Multivariate Normalverteilung

# Modellformulierung
n_1    = 40                                   # Anzahl von Datenpunkten Gruppe 1
n_2    = 40                                   # Anzahl von Datenpunkten Gruppe 2
n      = n_1 + n_2                            # Gesamtanzahl Datenpunkte
p      = 2                                    # Anzahl von Betaparameter
X      = matrix(c(rep(1,n_1), rep(0,n_1),     # Designmatrix
                  rep(0,n_2), rep(1,n_2)),
                  nrow  = n)
I_n    = diag(n)                              # n x n Einheitsmatrix
beta   = matrix(c(1,2), nrow = p)             # wahrer, aber unbekannter, Betaparameter
sigsqr = 14                                   # wahrer, aber unbekannter, Varianzparameter

# Datenrealisierung
y      = mvrnorm(1, X %*% beta, sigsqr*I_n)   # eine Realisierung eines n-dimensionalen ZVs
```

Theorem (Parameterschätzung im Zweistichproben-T-Test-Modell)

Gegeben sei die Designmatrixform des Zweistichproben-T-Test-Modells. Dann ergeben sich für den Betaparameterschätzer

$$\hat{\beta} = \begin{pmatrix} \frac{1}{n_1} \sum_{j=1}^{n_1} y_{1j} \\ \frac{1}{n_2} \sum_{j=1}^{n_2} y_{2j} \end{pmatrix} =: \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \tag{20}$$

und für den Varianzparameterschätzer

$$\hat{\sigma}^2 = \frac{\sum_{j=1}^{n_1} (y_{1j} - \bar{y}_1)^2 + \sum_{j=1}^{n_2} (y_{2j} - \bar{y}_2)^2}{n_1 + n_2 - 2} =: s_{12}^2 \tag{21}$$

Bemerkungen

- $\bar{y}_1$ und $\bar{y}_2$ bezeichnen die gruppenspezifischen Stichprobenmittel.
- s_{12}^2 wird als *gepoolte Stichprobenvarianz* bezeichnet.

Beweis

Für $i = 1, 2$ sei $y_i := (y_{i1}, ..., y_{in_i})^T$. Dann ergibt sich für den Betaparameterschätzer

$$
\begin{aligned}
\hat{\beta} &= (X^T X)^{-1} X^T y \\
&= \left(\begin{pmatrix} 1_{n_1} & 0_{n_2} \\ 0_{n_1} & 1_{n_2} \end{pmatrix} \begin{pmatrix} 1_{n_1} & 0_{n_1} \\ 0_{n_2} & 1_{n_2} \end{pmatrix} \right)^{-1} \begin{pmatrix} 1_{n_1} & 0_{n_2} \\ 0_{n_1} & 1_{n_2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= \begin{pmatrix} n_1 & 0 \\ 0 & n_2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{j=1}^{n_1} y_{1j} \\ \sum_{j=1}^{n_2} y_{2j} \end{pmatrix} \\
&= \begin{pmatrix} n_1^{-1} & 0 \\ 0 & n_2^{-1} \end{pmatrix} \begin{pmatrix} \sum_{j=1}^{n_1} y_{1j} \\ \sum_{j=1}^{n_2} y_{2j} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{n_1} \sum_{j=1}^{n_1} y_{1j} \\ \frac{1}{n_2} \sum_{j=1}^{n_2} y_{2j} \end{pmatrix} \\
&=: \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix}.
\end{aligned} \tag{22}
$$

Beweis (fortgeführt)

Gleichsam ergibt sich für Varianzparameterschätzer mit $n = n_1 + n_2$ und $p = 2$

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{n - p} \\
&= \frac{1}{n_1 + n_2 - 2} \left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} 1_{n_1} & 0_{n_1} \\ 0_{n_2} & 1_{n_2} \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right)^T \left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} 1_{n_1} & 0_{n_1} \\ 0_{n_2} & 1_{n_2} \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) \\
&= \frac{1}{n_1 + n_2 - 2} \begin{pmatrix} y_{11} - \bar{y}_1 \\ \vdots \\ y_{1n_1} - \bar{y}_1 \\ y_{21} - \bar{y}_2 \\ \vdots \\ y_{2n_2} - \bar{y}_2 \end{pmatrix}^T \begin{pmatrix} y_{11} - \bar{y}_1 \\ \vdots \\ y_{1n_1} - \bar{y}_1 \\ y_{21} - \bar{y}_2 \\ \vdots \\ y_{2n_2} - \bar{y}_2 \end{pmatrix} \\
&= \frac{\sum_{j=1}^{n_1} (y_{1j} - \bar{y}_1)^2 + \sum_{j=1}^{n_2} (y_{2j} - \bar{y}_2)^2}{n_1 + n_2 - 2} \\
&=: s_{12}^2 .
\end{aligned}
\tag{23}
$$

```r
# Dateneinlesen
fname      = file.path(getwd(), "5_Daten", "Zwei_Gruppen_Design.csv")  # Dateiname
D          = read.table(fname, sep = ",", header = TRUE)                # Dataframe
y_1        = D$dBDI[D$Group == "T"]                                     # BDI Differenzwerte in der F2F Gruppe
y_2        = D$dBDI[D$Group == "C"]                                     # BDI Differenzwerte in der ONL Gruppe

# Modellformulierung
n_1        = length(y_1)                                                # Anzahl Datenpunkte Gruppe 1 (F2F)
n_2        = length(y_1)                                                # Anzahl Datenpunkte Gruppe 2 (ONL)
n          = n_1 + n_2                                                  # Gesamtanzahl Datenpunkte
y          = matrix(c(y_1, y_2), nrow = n)                              # Datenvektor
p          = 2                                                          # Anzahl Betaparameter
X          = matrix(c(rep(1,n_1), rep(0,n_1),                           # Designmatrix
                      rep(0,n_2), rep(1,n_2)),
                      nrow  = n)

# Modellschätzung
beta_hat   = solve(t(X) %*% X) %*% t(X) %*% y                          # Betaparameterschätzer
eps_hat    = y - X %*% beta_hat                                         # Residuenvektor
sigsqr_hat = (t(eps_hat) %*% eps_hat) /(n-p)                            # Varianzparameterschätzer
s_sqr_12   = ((n_1-1)*var(y_1) + (n_2-1)*var(y_2))/(n_1+n_2-2)          # gepoolte Stichprobenvarianz
```

```
hat{beta}           :  3.067 -0.4
bar{y}_1, bar{y_2}  :  3.067 -0.4
hat{sigsqr}         :  14.59
s_12^2              :  14.59
```

Theorem (T-Teststatistik des Zweistichproben-T-Tests)

Gegeben sei die Designmatrixform des Zweistichproben-T-Tests. Dann ergibt sich für die T-Teststatistik mit

$$c := (1, -1)^T \text{ und } c^T\beta_0 =: \mu_0, \tag{24}$$

dass

$$T = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\left(\frac{\bar{y}_1 - \bar{y}_2 - \mu_0}{s_{12}}\right) \tag{25}$$

und es gilt

$$T \sim t(\delta, n_1 + n_2 - 2) \text{ mit } \delta = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\left(\frac{\mu_1 - \mu_2 - \mu_0}{\sigma}\right). \tag{26}$$

Bemerkungen

- Das Theorem basiert auf dem T-Statistik Theorem in Einheit (7) T-Statistiken.
- Wir erinnern an *Cohen's* d beim Zwei-Gruppen-Design

$$d := \frac{\bar{y_1} - \bar{y_2}}{s_{12}}. \tag{27}$$

- Offenbar gilt für dieses *Cohen's* d, dass mit $\mu_0 := 0$

$$T = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\, d \Leftrightarrow d = T/\sqrt{\frac{n_1 n_2}{n_1 + n_2}}. \tag{28}$$

Beweis

Mit dem T-Statistik Theorem gilt zunächst für die Zähler von T und δ, dass

$$c^T\hat{\beta} - c^T\beta_0 = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} - \mu_0 = \bar{y}_1 - \bar{y}_2 - \mu_0 \tag{29}$$

und

$$c^T\beta - c^T\beta_0 = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} - \mu_0 = \mu_1 - \mu_2 - \mu_0, \tag{30}$$

respektive. Weiterhin gilt für die Nenner von T und δ, dass

$$c^T(X^TX)^{-1}c = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} n_1^{-1} & 0 \\ 0 & n_2^{-1} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} n_1^{-1} & -n_2^{-1} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{n_1} + \frac{1}{n_2} \tag{31}$$

Außerdem gilt

$$\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-\frac{1}{2}} = \left(\frac{n_2}{n_1n_2} + \frac{n_1}{n_1n_2}\right)^{-\frac{1}{2}} = \left(\frac{n_1+n_2}{n_1n_2}\right)^{-\frac{1}{2}} = \left(\frac{n_1n_2}{n_1+n_2}\right)^{\frac{1}{2}} \tag{32}$$

Zusammengenommen folgt direkt, dass

$$T = \sqrt{\frac{n_1n_2}{n_1+n_2}}\left(\frac{\bar{y}_1 - \bar{y}_2 - \mu_0}{s_{12}}\right) \text{ und } \delta = \sqrt{\frac{n_1n_2}{n_1+n_2}}\left(\frac{\mu_1 - \mu_2 - \mu_0}{\sigma}\right). \tag{33}$$

Definition (Zweiseitiger Zweistichproben-T-Test)

Gegeben sei das Zweistichproben-T-Test-Modell. Für ein $\mu_0 \in \mathbb{R}$ seien die Nullhypothese und die Alternativhypothese gegeben durch

$$H_0 : \mu_1 - \mu_2 = \mu_0 \Leftrightarrow \Theta_0 := \{(\mu_1, \mu_2) \in \mathbb{R}^2 | \mu_1 - \mu_2 = \mu_0\} \tag{34}$$

und

$$H_1 : \mu_1 - \mu_2 \neq \mu_0 \Leftrightarrow \Theta_1 := \{(\mu_1, \mu_2) \in \mathbb{R}^2 | \mu_1 - \mu_2 \neq \mu_0\}, \tag{35}$$

respektive. Weiterhin sei die T-Teststatistik definiert durch

$$T := \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\bar{y}_1 - \bar{y}_2 - \mu_0}{s_{12}} \right) \tag{36}$$

Dann ist der zweiseitige *Zweistichproben-T-Teststatistik* definiert als der kritischen Wert-basierte Test

$$\phi(y) := 1_{\{|T| \geq k\}} = \begin{cases} 1 & |T| \geq k \\ 0 & |T| < k \end{cases}. \tag{37}$$

Bemerkungen

- Ausführlicher handelt es sich um den *zweiseitigen Zweistichproben-T-Test mit ungerichteter Hypothese*.

Allgemeines Lineares Modell

Linear Mixed Models

Selbstkontrollfragen

Definition (Linear Mixed Model)

Es sei

$$y = X_f\beta_f + X_r\beta_r + \varepsilon, \tag{38}$$

wobei

- y ein n-dimensionaler beobachtbarer Zufallsvektor ist, der *Daten* genannt wird,
- $X_f \in \mathbb{R}^{n\times p}$ eine vorgegeben Matrix ist, die *Fixed-Effects-Designmatrix* genannt wird,
- $\beta_f \in \mathbb{R}^p$ ein unbekannter fester Vektor ist, der *Fixed Effects* genannt wird,
- $X_r \in \mathbb{R}^{n\times q}$ eine vorgegeben Matrix ist, die *Random-Effects-Designmatrix* genannt wird,
- β_r ein q-dimensionaler latenter Zufallsvektor ist der *Random Effects* genannt wird und für den gilt, dass

$$\beta_r \sim N(\mu_{\beta_r}, \Sigma_{\beta_r}) \text{ mit } \mu_{\beta_r} \in \mathbb{R}^q \text{ und } \Sigma_{\beta_r} \in \mathbb{R}^{q\times q} \text{ p.d.,} \tag{39}$$

- ε ein n-dimensionaler latenter Zufallsvektor ist, der *Zufallsfehler* genannt wird und für den gilt, dass

$$\varepsilon \sim N(0_n, \Sigma_\varepsilon) \text{ mit } \Sigma_\varepsilon \in \mathbb{R}^{n\times n} \text{ p.d. und unabhängig von } \beta_r. \tag{40}$$

Dann heißt (38) *Linear Mixed Model*.

Bemerkungen

- Häufig gelten $\mu_{\beta_r} := 0_q, \Sigma_{\beta_r} := \sigma^2_{\beta_r} I_q, \sigma^2_{\beta_r} > 0$ und $\Sigma_\varepsilon := \sigma^2_\varepsilon I_n, \sigma^2_\varepsilon > 0$.
- Man mag die Darstellung des Linear Mixed Models in dieser Definition auch als *strukturelle Form* bezeichnen.

Definition (Verteilungsdarstellung des Linear Mixed Models)

Gegeben sei ein Linear Mixed Model

$$y = X_f\beta_f + X_r\beta_r + \varepsilon \text{ mit } \beta_r \sim N(\mu_{\beta_r}, \Sigma_{\beta_r}) \text{ und } \varepsilon \sim N(0_n, \Sigma_\varepsilon). \tag{41}$$

Dann nennt man die äquivalente Darstellung dieses Modells mit der marginalen Verteilung

$$\beta_r \sim N(\mu_{\beta_r}, \Sigma_{\beta_r}) \tag{42}$$

und der bedingten Verteilung

$$y \,|\, \beta_r \sim N(X_f\beta_f + X_r\beta_r, \Sigma_\varepsilon) \tag{43}$$

die Verteilungsdarstellung des Linear Mixed Models

Bemerkungen

- Die Äquivalenz folgt mit dem Theorem zur linear-affinen Transformation multivariate Normalverteilungen.
- Intuitiv beschreibt der Ausdruck $y = X_f\beta_f + X_r\beta_r + \varepsilon$ vor allem eine bedingte Verteilung.
- Die Fehlerkovarianzmatrix Σ_ε ist die Kovarianzmatrix dieser bedingten Verteilung.

Theorem (Gemeinsame Verteilung des Linear Mixed Models)

Gegeben sei ein Linear Mixed Model. Dann gilt für die gemeinsame Verteilung von Daten und Random Effects, dass

$$\begin{pmatrix} \beta_r \\ y \end{pmatrix} \sim N\left(\mu_{\beta_r,y}, \Sigma_{\beta_r,y}\right) \tag{44}$$

mit

$$\mu_{\beta_r,y} := \begin{pmatrix} \mu_{\beta_r} \\ X_f\beta_f + X_r\mu_{\beta_r} \end{pmatrix} \in \mathbb{R}^{q+n} \text{ und } \Sigma_{\beta_r,y} := \begin{pmatrix} \Sigma_{\beta_r} & \Sigma_{\beta_r}X_r^T \\ X_r\Sigma_{\beta_r} & X_r\Sigma_{\beta_r}X_r^T + \Sigma_\varepsilon \end{pmatrix} \in \mathbb{R}^{(q+n)\times(q+n)} \tag{45}$$

Beweis

Die gemeinsame Verteilung des Linear Mixed Models ergibt sich direkt durch Anwendung des Theorems zu Gemeinsamen Normalverteilungen auf die Verteilungsdarstellung des Linear Mixed Models. □

Theorem (Marginale Datenverteilung des Linear Mixed Models)

Gegeben sei ein Linear Mixed Model. Dann gilt für die marginale Verteilung der Daten, dass

$$y \sim N\left(\mu_y, \Sigma_y\right) \tag{46}$$

mit

$$\mu_y := X_f\beta_f + X_r\mu_{\beta_r} \in \mathbb{R}^n \text{ und } \Sigma_y := X_r\Sigma_{\beta_r}X_r^T + \Sigma_\varepsilon \in \mathbb{R}^{n\times n} \tag{47}$$

Beweis

Die Aussage ergibt sich direkt aus dem Theorem zur Gemeinsamen Verteilung des Linear Mixed Models und dem Theorem zu Marginalen Normalverteilungen. □

Bemerkungen

- Linear Mixed Models erlauben es, nicht-sphärische Kovarianzmatrixstrukturen zu modellieren.
- Gilt speziell $\Sigma_{\beta_r} := \sigma^2_{\beta_r} I_q, \sigma^2_{\beta_r} > 0$ und $\Sigma_\varepsilon := \sigma^2_\varepsilon I_n, \sigma^2_\varepsilon > 0$, so folgt

$$y \sim N\left(X_f\beta_f + X_r\mu_{\beta_r}, \sigma^2_{\beta_r} X_r X_r^T + \sigma^2_\varepsilon I_n\right) \tag{48}$$

- Parameter wie $\sigma^2_{\beta_r}$ und σ^2_ε nennt man *Kovarianzkomponenten*.

Definition (Hierarchische Darstellung des Linear Mixed Models)

Gegeben sei ein Linear Mixed Model

$$y = X_f\beta_f + X_r\beta_r + \varepsilon \text{ mit } \beta_r \sim N(\mu_{\beta_r}, \Sigma_{\beta_r}) \text{ und } \varepsilon \sim N(0_n, \Sigma_\varepsilon). \tag{49}$$

Dann nennt man die äquivalente Darstellung dieses Modells in der Form

$$\begin{aligned} \beta_r &= \mu_{\beta_r} + \eta && \text{mit } \eta \sim N(0_q, \Sigma_{\beta_r}) \Leftrightarrow \beta_r \sim N(\mu_{\beta_r}, \Sigma_{\beta_r}) \\ y &= X_f\beta_f + X_r\beta_r + \varepsilon && \text{mit } \varepsilon \sim N(0_n, \Sigma_\varepsilon) \Leftrightarrow y \,|\, \beta_r \sim N(X_f\beta_f + X_r\beta_r, \Sigma_\varepsilon) \end{aligned} \tag{50}$$

die *hierarchische Darstellung des Linear Mixed Models*

Bemerkung

- Man nennt diese Darstellung auch ein *Mehrebenenmodell*.
- Es ist leicht, sich Linear Mixed Models mit mehr als den hier spezifizierten zwei Ebenen vorzustellen.
- Die obigen Verteilungsaussagen gelten natürlich auch für die Hierarchische Form des Linear Mixed Models.

Beispiel (1) Allgemeines Lineares Modell

Das Allgemeine Linear Modell ist das Linear Mixed Modell

$$y = X_f\beta_f + X_r\beta_r + \varepsilon \text{ mit } \varepsilon \sim N\left(0_n, \Sigma_\varepsilon\right), \tag{51}$$

für das

- y den n-dimensionalen Datenvektor bezeichnet,
- $X_f := X \in \mathbb{R}^{n\times p}$ die Designmatrix bezeichnet,
- $\beta_f := \beta \in \mathbb{R}^p$ der wahre, aber unbekannte, Betaparametervektor ist,
- $X_r\beta_r := 0_n$ der Nullvektor ist bzw. nicht existiert
- ε der Zufallsfehler ist, für den angenommen wird, dass $\Sigma_\varepsilon := \sigma^2 I_n$.

Das ALM ist also das LMM ohne Random-Effects und sphärischem Kovarianzmatrixparameter.

Beispiel (2) Fixed-Effects-Modell der Metaanalyse

Das Fixed-Effects-Modell der Metaanalyse ist das Linear Mixed Model

$$y = X_f\beta_f + X_r\beta_r + \varepsilon \text{ mit } \varepsilon \sim N\left(0_n, \Sigma_\varepsilon\right), \tag{52}$$

für das

- y der Vektor von n studienspezifischen Effektstärkeschätzern Cohen's d_i oder Hedges' $g_i, i = 1, ..., n$ ist
- $X_f := 1_n \in \mathbb{R}^{n\times 1}$ die Fixed-Effects-Designmatrix bezeichnet,
- $\beta_f := \delta$ die wahre, aber unbekannte, Effektstärke bezeichnet
- $X_r\beta_r := 0_n$ der Nullvektor ist bzw. nicht existiert
- ε der Zufallsfehler ist, für den angenommen wird, dass

$$\Sigma_\varepsilon := \text{diag}(\sigma_1^2, ..., \sigma_n^2) \tag{53}$$

  mit bekannten studienspezifischen Varianzschätzern $\sigma_i^2, i = 1, ..., n$ ist.

Das Fixed-Effects-Modell der Metaanalyse ist also das LMM ohne Random-Effects bzw. das ALM mit einem eindimensionalen Betaparametervektor und nicht-sphärischem, diagonalem und bekanntem Kovarianzmatrixparameter.

Beispiel (3) Random-Effects-Modell der Metaanalyse

Das Random-Effects-Modell der Metaanalyse ist das Linear Mixed Model

$$y = X_f\beta_f + X_r\beta_r + \varepsilon \text{ mit } \varepsilon \sim N\left(0_n, \Sigma_\varepsilon\right), \quad (54)$$

für das

- y der Vektor von n studienspezifischen Effektstärkeschätzern Cohen's d_i oder Hedges' $g_i, i = 1, ..., n$ ist,
- $X_f\beta_f := 0_n$ der Nullvektor ist bzw. nicht existiert,
- $X_r := I_n$ die $n \times n$-dimensionale Einheitsmatrix ist,
- $\beta_r := (\delta_1, ..., \delta_n)^T$ der Vektor der latenten standardisierten Effektstärken ist, für den angenommen wird, dass

  $$\beta_r \sim N\left(1_n\delta, \tau^2 I_n\right), \quad (55)$$

  wobei δ die wahre, aber unbekannte, standardisierte Effektstärke und τ^2 die wahre, aber unbekannte, studienübergreifende Effektstärkevarianz bezeichnen und
- ε der Zufallsfehler ist, für den angenommen wird, dass

  $$\Sigma_\varepsilon := \text{diag}(\sigma_1^2, ..., \sigma_n^2) \quad (56)$$

  mit bekannten studienspezifischen Varianzschätzern $\sigma_i^2, i = 1, ..., n$ ist.

Das Random-Effects-Modell der Metaanalyse ist also das LMM ohne Fixed-Effects und mit der Annahme, dass jedem studienspezifischen Effektstärkeschätzer y_i ein studienspezifischer Random-Effect δ_i zugeordnet werden kann, welcher wiederrum eine normalverteilte Zufallsvariable mit Erwartungswertparameter δ ist und schließlich für den Zufallsfehlers der studienspezifischen Effektstärkeschätzer y_i ein nicht-sphärischer, diagonaler und bekannter Kovarianzmatrixparameter angenommen wird.

Allgemeines Lineares Modell

Linear Mixed Models

Selbstkontrollfragen

Selbstkontrollfragen

1. Geben Sie die Definition des Linear Mixed Models wieder.
2. Geben Sie die Definition der Verteilungsdarstellung des Linear Mixed Models wieder.
3. Geben Sie das Theorem zur Gemeinsamen Verteilung des Linear Mixed Models wieder.
4. Geben Sie das Theorem zur Marginalen Datenverteilung des Linear Mixed Models wieder.
5. Geben Sie die Definition der Hierarchischen Darstellung des Linear Mixed Models wieder.
6. Formulieren Sie das Allgemeine Lineare Model als Linear Mixed Model.
7. Formulieren Sie das Fixed-Effects-Modell der Metaanalyse als Linear Mixed Model.
8. Formulieren Sie das Random-Effects-Modell der Metaanalyse als Linear Mixed Model.